

**Translation of S64-44498**

Fig. 2 illustrates a specific example of speech unit file employed in the present invention, particularly as the speech unit set including each top-ranking 50 phonological junctures (number of sound junctures: 1-4) in occurrence frequency for the entry words in the language dictionary as its subset. In Fig. 2, numeral figures shown on the right of each phonological juncture represent occurrence frequencies, which are listed in order of high to low frequency respectively for each number of junctures. Also in Fig. 2, 108 phonological junctures enclosed by circles comprise of a unit file, which cover each top-ranking 50 phonological junctures of 1 through 4 junctures, together by their sub combinations of the junctures.

To describe more specifically, the phonological sound of 1 juncture, "u", is included into the phonological sound of 2 junctures, "ru", which is also included into the phonological sound of 3 junctures, "eru". This is also included into the phonological sound of 4 junctures "keru". Therefore, the phonological sound "u", "ru" and "eru" can be obtained from the phonological sound "keru" as its sub-unit without storing them in the speech unit file.

In Fig. 3, the example shown in Fig. 2 is used as a speech unit set to illustrate an example of a series of speech units selected, for example, when /sakuragasaku/ is entered.

As shown in Fig. 3, the beginning of the entry, /sa/, is obtained from a part of the phonological sound of 4 junctures, /saku/. Also /kura/ is obtained from the phonological sound of 4 junctures, /kura/. This results in a series of speech units of "sakura". In addition, "ga" is obtained from /ga/ in the phonological sound of 4 junctures, /gaku/, as well as "saku" from the phonological sound of 4 junctures, /saku/.



Relation between a length of speech synthesis unit and problems of the speech synthesis system construction.

Problems of the speech synthesis system construction	Length of speech synthesis unit	
	Short	Long
Minimum storage capacity required for storing units	Small	Large
Number of units, and manpower required for producing speech units	Low	High
Availability and flexibility for overlapping usage of units	High	Low
Necessity of unit deformation rules due to phonological environment	High	Low
Necessity of connection rules between units	High	Low

(The terms "High", "Low" and "Small" enclosed by circles are preferred for the system construction.)

# VOICE SYNTHESIZING SYSTEM USING COMPOSITE VOICE UNIT

3

Patent Number: JP1044498  
Publication date: 1989-02-16  
Inventor(s): KOSAKA YOSHINORI  
Applicant(s): A T R JIDO HONYAKU DENWA KENKYUSHO:KK  
Requested Patent:  JP1044498  
Application Number: JP19870202286 19870812  
Priority Number(s):  
IPC Classification: G10L5/02  
EC Classification:  
Equivalents: JP2894447B2

## Abstract

**PURPOSE:** To regulate a unit set corresponding to a preparing condition and to improve the sound quality by executing voice synthesis by using a composite voice unit having various stored length and configuration and partial units included in the composite voice unit.

**CONSTITUTION:** A rhythm control signal is inputted from an input terminal 1 to a voice unit selection part 2. The selection part 2 selects voice units matched with the unit signal synthesis of input signals and selects also a concrete constitution condition suitable for a synthetic condition from a voice control knowledge file 4. The voice units selected by the selection part 2 are combined by a unit combining part 5 and the combined unit is inputted to a synthetic waveform generating part 6. The generating part 6 generates a synthetic wave based upon an inputted unit sequence and outputs the generated synthetic waveform from an output terminal 7. In the case of selecting voice units having various structure and units, the same synthesizer can be used. In addition, sound quality can be improved by selecting a suitable synthetic condition.

Data supplied from the **esp@cenet** database - I2

⑯ 日本国特許庁 (JP)

⑯ 特許出願公開

## ⑯ 公開特許公報 (A)

昭64-44498

⑯ Int.CI.<sup>4</sup>

G 10 L 5/02

識別記号

府内整理番号

K-8622-5D

⑯ 公開 昭和64年(1989)2月16日

審査請求 未請求 発明の数 1 (全5頁)

⑯ 発明の名称 複合音声単位を用いた音声合成方式

⑯ 特 願 昭62-202286

⑯ 出 願 昭62(1987)8月12日

⑯ 発明者 勾坂 芳典 京都府相楽郡精華町大字乾谷小字三平谷5番地 株式会社  
エイ・ティ・アール自動翻訳電話研究所内

⑯ 出願人 株式会社 エイ・ティ 京都府相楽郡精華町大字乾谷小字三平谷5番地  
エイ・ティ・アール自動翻訳電話研究所

⑯ 代理人 弁理士 深見 久郎 外2名

## 明細書

## 1. 発明の名称

複合音声単位を用いた音声合成方式

## 2. 特許請求の範囲

音声単位の編集により音声合成を行なうかあるいは規則により任意語彙の音声合成を行なう方式において、

音素レベルから音節、形態素、単語、文節さらに文章に至るまでの種々の長さ、構造を有する複合音声単位を蓄え、それらの複合音声単位ならびにそれらに含まれる部分単位を用いて音声合成を行なうことを特徴とする、複合音声単位を用いた音声合成方式。

## 3. 発明の詳細な説明

## [産業上の利用分野]

この発明は複合音声単位を用いた音声合成方式に関し、特に、音声単位の編集、結合、変形によって音声合成を行なうような音声合成方式に関する。

## [従来の技術]

従来の編集型音声合成方式、規則による音声合成方式において種々の音声単位が提案されている。このような音声単位には、短いものから順に、音素、CV(子音、母音)音節、半音節(子音、母音連節および母音・子音連接、VCV(母音・子音・母音)連接、CVC(子音・母音・子音)連接、CVCV(子音・母音・子音・母音)連接、単語、文節といったものがある。

従来の音声合成方式では、これらのうちのどれか1種類の音声単位集合を基本とし、他の1~2種類の音声単位を補助的に使用する場合を含めても、たかだか2~3種類の音声単位の編集、結合、変形により音声を生成していた。

## [発明が解決しようとする問題点]

第1図は従来の各種の音声単位を単位長さの点で短いものと長いものの2種類に大別し、音声合成システム構成上の優劣を比較した図である。第1図において、音声合成システムを構成する上で問題となる単位の蓄積に必要な記憶容量は少ない方が望ましく、音声単位の長さが短いほど有利である。

ある。また、合成単位を作成する工数は同じく音声単位の長さが短いほど有利である。単位重複の利用可能性や柔軟性といった面からすれば、音声単位が短いほど大きくなる。一方、音声品質を左右する単位変形法や単位接続法などにおいては、音声単位の長さが長いほど問題点を少なくすることができます。上述の対比から明らかのように、2種類の音声単位だけでバランス良く音声合成システムを構成することは極めて困難であった。

それゆえに、この発明の主たる目的は、種々の構造を持つ多種類の音声単位とそれらに含まれる部分単位を使用することにより、より効率の良い音声合成を得ることのできるような複合音声単位を用いた音声合成方式を提供することである。

【問題点を解決するための手段】

この発明は複合音声単位を用いた音声合成方式であって、音楽レベルから音節、形態素、単語、文節さらに文章に至るまでの種々の長さ、構造を有する複合音声単位を蓄えておき、それらの複合音声単位ならびにそれらに含まれる部分単位を用

- 3 -

そして、音声単位選択部2は、音声単位列の選択に際して、音声制御知識ファイル4から、当該合成箇所の音韻環境、韻律情報を基にして、望ましい合成単位の抽出環境、基本周波数、音韻時間長などの具体的構成条件が与えられ、これらの合成条件を単位選択の基準として用いる。音声単位選択部2で選択された音声単位列は単位結合部5に与えられ、選択された単位どうしの結合が行なわれる。結合された音声単位は合成波形生成部6に与えられ、単位結合部5で得られた合成単位系列を基にして合成波形を生成する。そして、生成された音声は出力端子7に出力される。

また、この発明による方式は、種々の構造を持つ多種類の単位を使用する点に特色があり、この発明による方式を実施する際に用いられる音声単位を表現する音響パラメータならびに単位選択に用いる種々の基準、音声制御知識、音声波形を生成するための合成器については、何らこれを規定するものではなく、すべてに適用可能である。

第2図はこの発明に用いられる音声単位ファイ

- 5 -

いて音声合成を行なうように構成したものである。

【作用】

この発明に係る複合音声単位を用いた音声合成方式は、蓄えた複合音声単位およびそれらに含まれる部分単位を用いて音声合成を行なう。

【発明の実施例】

以下に、図面を参照してこの発明の実施例について詳細に説明する。

第1図はこの発明の音声合成方式の概略ブロック図である。第1図において、入力端子1には、出力したい音声内容に対応した音韻系列信号、アクセント、息継ぎなどの韻律制御信号が入力される。入力端子1に入力された韻律制御信号は音声単位選択部2に与えられる。この音声単位選択部2は、入力情報を基にして、音声単位ファイル3に予め蓄えられた種々の構造を持つ多数の音声単位から合成に適した音声単位列を選択する。音声制御知識ファイル4には、抽出環境、基本周波数、音韻時間長などの具体的合成条件を与える規則、諸表が蓄えられている。

- 4 -

ルの具体的な一例を示す図であり、特に、国語辞書内の見出語中の音韻連接（連接数1～4）のうち、高頻度のもの上位50個を、その部分集合として含む音声単位セットとして示したものである。第2図において、各音韻連接の右の数字は出現頻度を示し、各音韻連接数ごとに頻度の高い順から並べてある。また、第2図において、丸印で囲んだ音韻連接108個が単位ファイルを構成し、それらの部分連接により、これら1～4連接上位50個を包含している。

より具体的に説明すると、連接数1の音韻“u”は連接数2の音韻“ru”に含まれており、この音韻“ru”は連接数3の音韻“eru”に含まれており、この音韻“eru”は連接数4の音韻“keru”に含まれている。したがって、音韻“u”、“ru”、“eru”については、音声単位ファイルに蓄えておく必要がなく、音韻“keru”から部分単位として抽出することによって得られることになる。

第3図は第2図で示した例を音声単位セットと

- 6 -

して使用し、たとえば「桜が咲く」／s a k u r a g a s a k u／を入力した場合に選択される音声単位系列の一例を示す図である。

第3図に示すように、文頭の／s a／は4連接／s a k u／から生成される部分列によって与えられ、／k u r a／は4連接／k u r a／によって与えられ、「桜」の音声単位列が実現される。さらに、「が」は4連接／g a k u／中の／g a／によって実現され、「咲く」は4連接／s a k u／によって同様に実現される。

#### [発明の効果]

以上のように、この発明によれば、音声単位の多様化により、単位蓄積に許容される記憶容量、作成に許容される工数など作成上の要求条件に応じて、自由に効率良く単位セットを規定できる利点を有するとともに、長い音声単位中に含まれる短い連接の部分利用など、単位を有効かつ柔軟に使用できる利点を有している。さらに、より長い単位を利用するため、単位間の接続に起因する音声品質の劣化が少なく、良好な音声品質を期待

できる。

#### 4. 図面の簡単な説明

第1図はこの発明の方式を実施するための概略ブロック図である。第2図は音声単位ファイルを構成する具体的な一例を示す図である。第3図は入力音韻系列に対して選択される音声単位系列の一例を示す図である。第4図は音声合成単位長と音声合成システム構成上の問題点の関係を示す図である。

図において、1は入力端子、2は音声単位選択部、3は音声単位ファイル、4は音声制御知識ファイル、5は単位結合部、6は合成波形生成部、7は出力端子を示す。

特許出願人 株式会社エイ・ティ・アール

自動翻訳電話研究所

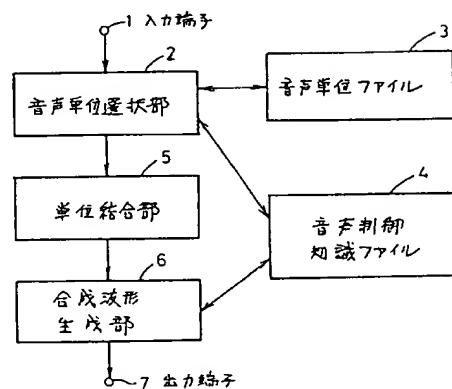
代理人 弁理士 深見久郎  
(ほか2名)



- 7 -

- 8 -

第1図



連接名	古語名	出現頻度	古語連接名	出現頻度	第2図			
					音韻連接名	出現頻度	音韻連接名	出現頻度
1	u	53264	ou	12394	au	3005	shou	355
2	i	44069	ku	8675	oku	2759	kaku	350
3	a	43367	ka	7005	kou	1782	oushi	347
4	o	36832	shi	6389	eru	1516	seki	337
5	ɔ	28800	ak	5558	sei	1234	oka	336
6	e	21869	ki	5462	kal	216	mono	322
7	ɛ	17702	ai	5445	ash	1205	koku	297
8	ɛ	15749	aw	5241	shou	1202	esk	297
9	s	12681	ch	5194	shin	1094	shiu	290
10	ɛ	11061	tsu	4866	ari	1083	fuku	283
11	ɛ	10795	ko	4766	al	1067	renu	274
12	ɛ	10214	ri	4268	ar	1065	tori	272
13	ɛ	7416	ru	4214	as	1039	keru	272
14	ɛ	7309	ox	4085	eki	1030	setsu	268
15	ə	65777	ei	3935	uku	1006	saku	251
16	ə	6352	lo	3921	ush	990	usei	250
17	ə	6391	ta	3915	ir	977	ouse	246
18	ə	4334	uk	3903	ki	916	kata	246
19	ə	4875	iw	3883	tsu	916	shiki	244
20	ə	4375	ar	3718	tsu	965	okuo	238
21	ə	4103	ar	3580	uka	949	suru	237
22	ə	3898	su	3575	ori	905	soku	227
23	ə	2183	ur	3563	ir	921	ukou	225
24	ə	1752	se	3542	ok	911	shou	217
25	ə	1682	ik	3506	seh	904	shitsu	215
26	ə	1441	sa	3230	sou	887	ito	213
27	ə	1153	ra	3163	aru	887	to	208
28	ə	951	ma	3034	ra	832	to	207
29	ə	591	ke	2887	tou	834	utan	199
30	ə	501	ha	2644	tsu	853	hi	198
31	ə	484	ra	2221	tsu	845	ouji	195
32	ə	380	ra	2396	tu	833	rau	193
33	ə	338	ss	352	ji	832	kusa	193
34	ə	34	hy	274	chi	821	ushin	189
35	ə	213	iv	239	ur	818	eru	185
36	ə	210	ey	239	ur	812	kuse	185
37	ə	164	or	2260	ur	802	uto	186
38	ə	149	or	2248	la	774	akus	186
39	ə	101	so	2243	ur	765	oku	186
40	ə	52	ia	2204	ush	730	kura	185
41	ə	34	sho	2191	ro	694	kata	183
42	ə	27	no	2148	kak	678	haku	183
43	ə	15	us	2123	oshi	676	ui	182
44	ə	14	re	2115	san	672	osu	180
45	ə	10	re	2071	uto	671	akush	179
46	ə	8	ok	2065	to	670	ui	177
47	ə	7	da	2040	chi	668	iseli	177
48	ə	5	da	2023	sh	665	iseli	177
49	ə	3	or	2012	us	653	akur	177

音声合成単位と音声合成システム構成上の問題点の関係	第4図		
	音声合成システム構成上の問題点	合成用音声単位の長さ	合成用音声単位の長さ
単位の音素に最低限必要な記憶容量	①	短	長
単位個数、単位音声作成に要する工数	①	大	大
単位の重複利用可能性、柔軟性	②	小	小
音韻環境による単位変形規則の必要性	大	①	①
単位間接接続規則の必要性	大	①	①

(音声合成単位と音声合成システム構成上の問題点の関係)

手 続 補 正 書

昭和63年2月10日

特許庁長官殿

1、事件の表示

昭和62年特許願第 202286 号

適

6. 補正の対象

明細書の発明の詳細な説明の欄

7. 補正の内容

明細書第2頁第15行および第17行ないし第18行の「第1図」を「第4図」に補正する。

以上

2、発明の名称

複合音声単位を用いた音声合成方式

3、補正をする者

事件との関係 特許出願人

住 所 京都府相楽郡精華町大字乾谷小字三平谷5番地

名 称 株式会社エイ・ティ・アール自動翻訳電話研究所

代表者 横松 明

4、代理 人

住 所 大阪市北区南森町2丁目1番29号 住友銀行南森町ビル

電話 大阪(06)361-2021(代)

氏 名 弁理士(6474) 深見久郎



5、補正命令の日付

自発補正

方 式  
箱 送



- 2 -